Paired and Unpaired Deep Generative Models on Multimodal Retinal Image Reconstruction [†]

Álvaro S. Hervella [1,2,*], José Rouco [1,2], Jorge Novo [1,2] and Marcos Ortega [1,2]

[1] CITIC-Research Center of Information and Communication Technologies, Universidade da Coruña, 17051 A Coruña, Spain

[2] Department of Computer Science, Universidade da Coruña, 17051 A Coruña, Spain

[*] Correspondence: a.suarezh@udc.es; Tel.: +34-881011307

[†] Presented at 2nd XoveTIC Conference, A Coruña, Spain, 5–6 September 2019.

check for updates

Abstract: This work explores the use of paired and unpaired data for training deep neural networks in the multimodal reconstruction of retinal images. Particularly, we focus on the reconstruction of fluorescein angiography from retinography, which are two complementary representations of the eye fundus. The performed experiments allow to compare the paired and unpaired alternatives.

1. Introduction

Recently, deep learning algorithms have made possible to estimate the mapping between different image modalities. In that regard, different methods have been proposed to learn the multimodal transformation using paired [1] or unpaired data [2]. In medical imaging, the paired data can be easily gathered due to the common use of complementary image modalities in the clinical practice. Moreover, the paired data represents a rich and complementary source of information that facilitates the training of a neural network. However, to successfully exploit the available paired data, the multimodal images must be previously registered. This multimodal registration is a challenging task that may fail in the most complex scenarios.

In this work, we explore the use of paired and unpaired data for training neural networks in the multimodal reconstruction of retinal images. Specifically, we target the generation of fluorescein angiography, which represents an invasive capture technique, from color retinography.

2. Methodology

For the paired scenario, we use the methodology proposed by Hervella et al. [1]. This method requires the registration of different image pairs, which is performed following the approach in [3]. Then, a generative neural network is trained using the Structural Similarity (SSIM) between the target image and the network output as loss function. For the unpaired case, we use the methodology known as CycleGAN [2], which employs a generative adversarial setting. In particular, a discriminator network learns to distinguish between real and generated angiographies. At the same time, an additional generator network learns to reconstruct the original retinographies from the generated angiographies. This enforces the structural coherence between input and generated images.

In order to produce a fair comparison, we use the same network architectures for both approaches [4]. The neural networks are trained using the Adam algorithm and data augmentation [4]. For the training and evaluation, we used 59 image pairs from the Isfahan MISP database and 59 additional pairs obtained from the Complexo Hospitalario Universitario de Santiago (CHUS). Half of the data is used for training and the other half is hold out as test set.

3. Results and Conclusions

Examples of generated angiographies from the test set are depicted in Figure 1. Both approaches allow to estimate an adequate transformation from retinography to angiography. However, the images generated using the paired method are more realistic due to the use of the adversarial setting. Regarding the reconstruction error in the test set, both approaches achieveD similar results by means of MAE or MSE. However, the paired method achieved a lower error by means of SSIM. This evidences that the images generated using the unpaired method are structurally less accurate. However, the implications of this result depend on the specific target application.

| (a) | (b) | (c) | (d) |
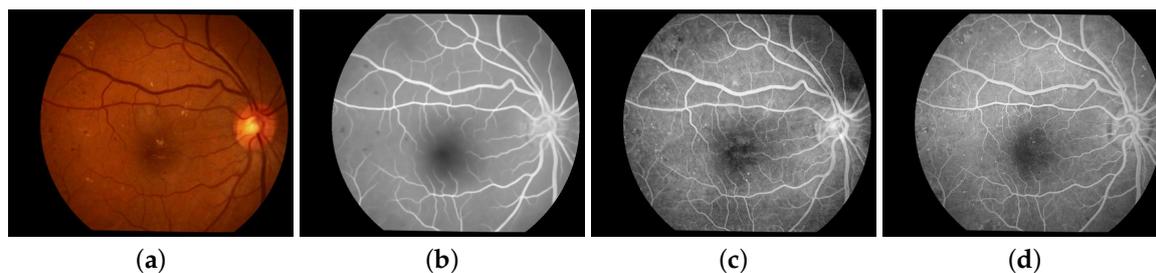
Figure 1. Examples of (**b**,**c**) generated angiographies together with (**a**) the original retinography and (**d**) the real angiography. (**b**) Generated using the paired approach. (**c**) Generated using the unpaired approach.

References

1. Hervella, A.S.; Rouco, J.; Novo, J.; Ortega, M. Retinal Image Understanding Emerges from Self-Supervised Multimodal Reconstruction. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI), Granada, Spain, 16–20 September 2018.
2. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.
3. Hervella, A.S.; Rouco, J.; Novo, J.; Ortega, M. Multimodal Registration of Retinal Images Using Domain-Specific Landmarks and Vessel Enhancement. In Proceedings of the International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES), Belgrade, Serbia, 3–5 September 2018.
4. Hervella, A.S.; Rouco, J.; Novo, J.; Ortega, M. Deep Multimodal Reconstruction of Retinal Images Using Paired or Unpaired Data. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019.